

10/530421

**RAW SEQUENCE LISTING**

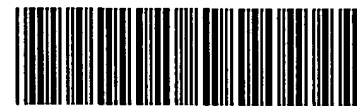
**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/530,421

Source: PG10

Date Processed by STIC: 11/8/05

***ENTERED***



PCT

## RAW SEQUENCE LISTING

DATE: 11/08/2005

PATENT APPLICATION: US/10/530,421

TIME: 12:07:48

Input Set : A:\54A-000170PC.ST251.txt

Output Set: N:\CRF4\11082005\J530421.raw

3 <110> APPLICANT: The Scripps Research Institute  
 4 Schultz, Peter  
 5 Wang, Lei  
 7 <120> TITLE OF INVENTION: SITE SPECIFIC INCORPORATION OF KETO AMINO ACIDS INTO  
 PROTEINS

9 <130> FILE REFERENCE: 54A-000170PC  
 C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/530,421  
 C--> 12 <141> CURRENT FILING DATE: 2005-04-05

14 <160> NUMBER OF SEQ ID NOS: 24  
 16 <170> SOFTWARE: PatentIn version 3.1

P.7

18 <210> SEQ ID NO: 1  
 19 <211> LENGTH: 918  
 20 <212> TYPE: DNA  
 21 <213> ORGANISM: artificial  
 23 <220> FEATURE:  
 24 <223> OTHER INFORMATION: tRNA synthetase based on M. jannaschii tyrosyl-tRNA  
 synthetase

26 <400> SEQUENCE: 1  
 27 atggacgaat ttgaaatgat aaagagaaac acatctgaaa ttatcagcga ggaagagtta 60  
 29 agagaggttt taaaaaaaga tgaaaaatct gctgggatag gttttgaacc aagtggtaaa 120  
 31 atacatttag ggcattatct ccaaataaaa aagatgattg atttcaaaa tgctggattt 180  
 33 gatataatta tattgttggc tgattttacac gcctatttaa accagaaagg agagttggat 240  
 35 gagattagaa aaataggaga ttataacaaa aaagtttttg aagcaatggg gttaaaggca 300  
 37 aaatgtgctt atggaagtcc tttccagctt gataaggatt atacactgaa tgtctataga 360  
 39 ttggctttta aaactacctt aaaaagagca agaaggagta tggaacttat agcaagagag 420  
 41 gatgaaaatc caaaggttgc tgaagttatc tatccaataa tgcagggttaa tggttatcat 480  
 43 tatcttggcg ttgatgttgc agttggaggg atggagcaga gaaaaataca catgttagca 540  
 45 agggagcttt taccaaaaaa gggtgtttgt attcacaacc ctgtcttaac gggtttggat 600  
 47 ggagaaggaa agatgagttc ttcaaaaggg aattttatag ctgttgatga ctctccagaa 660  
 49 gagattaggg ctaagataaa gaaagcatac tgcccagctg gagttgttga aggaaatcca 720  
 51 ataattggaga tagctaaata ctcccttgaa taccctttta ccataaaaag gccagaaaaa 780  
 53 tttggtggag atttgacagt taatagctat gaggagttag agagtttatt taaaaataag 840  
 55 gaattgcata caatggattt aaaaaatgct gtagctgaag aacttataaa gatttttagag 900  
 57 ccaattagaa agagatta 918

60 <210> SEQ ID NO: 2  
 61 <211> LENGTH: 918  
 62 <212> TYPE: DNA  
 63 <213> ORGANISM: artificial  
 65 <220> FEATURE:  
 66 <223> OTHER INFORMATION: tRNA synthetase based on M. jannaschii tyrosyl-tRNA  
 synthetase

68 <400> SEQUENCE: 2  
 69 atggacgaat ttgaaatgat aaagagaaac acatctgaaa ttatcagcga ggaagagtta 60  
 71 agagaggttt taaaaaaaga tgaaaaatct gctcagatag gttttgaacc aagtggtaaa 120

```
73 atacatttag ggcattatct ccaaataaaa aagatgattg atttacaaaa tgctggattt 180
75 gatataatta tattgttggc tgatttacac gctatttaa accagaaagg agagttggat 240
```

## RAW SEQUENCE LISTING

DATE: 11/08/2005

PATENT APPLICATION: US/10/530,421

TIME: 12:07:48

Input Set : A:\54A-000170PC.ST251.txt

Output Set: N:\CRF4\11082005\J530421.raw

```

77 gagattagaa aaataggaga ttataacaaa aaagtttttg aagcaatggg gttaaaggca 300
79 aaatatgttt atggaagtcc tttccagctt gataaggatt atacactgaa tgtctataga 360
81 ttggctttta aaactacctt aaaaagagca agaaggagta tggaacttat agcaagagag 420
83 gatgaaaatc caaagggttg tgaagttatc tatccaataa tgcagggtta ttgttctcat 480
85 tattatggcg ttgatgttgc agttggaggg atggagcaga gaaaaataca catgttagca 540
87 agggagcttt taccaaaaaa ggttgtttgt attcacaacc ctgtcttaac gggtttggat 600
89 ggagaaggaa agatgagttc ttcaaaaggg aattttatag ctggtgatga ctctccagaa 660
91 gagattaggg ctaagataaa gaaagcatac tgcccagctg gagttggtga aggaaatcca 720
93 ataattggaga tagctaaata cttccttgaa tctcctttaa ccataaaaag gccagaaaaa 780
95 tttggtggag atttgacagt taatagctat gaggagttag agagtttatt taaaaataag 840
97 gaattgcac caatggattt aaaaaatgct gtagctgaag aacttataaa gatttttagag 900
99 ccaattagaa agagatta 918

```

102 &lt;210&gt; SEQ ID NO: 3

103 &lt;211&gt; LENGTH: 918

104 &lt;212&gt; TYPE: DNA

105 &lt;213&gt; ORGANISM: artificial

107 &lt;220&gt; FEATURE:

108 &lt;223&gt; OTHER INFORMATION: tRNA synthetase based on M. jannaschii tyrosyl-tRNA synthetase

110 &lt;400&gt; SEQUENCE: 3

```

111 atggacgaat ttgaaatgat aaagagaaac acatctgaaa ttatcagcga ggaagagtta 60
113 agagagggttt taaaaaaaga tgaaaaatct gctactatag gttttgaacc aagtggtaaa 120
115 atacatttag ggcattatct ccaaataaaa aagatgattg atttcaaaa tgctggattt 180
117 gatataatta tattgttggc tgatttacac gcctatttaa accagaaagg agagttggat 240
119 gagattagaa aaataggaga ttataacaaa aaagtttttg aagcaatggg gttaaaggca 300
121 aaatatgttt atggaagtac gttccagctt gataaggatt atacactgaa tgtctataga 360
123 ttggctttta aaactacctt aaaaagagca agaaggagta tggaacttat agcaagagag 420
125 gatgaaaatc caaagggttg tgaagttatc tatccaataa tgcagggtta tccgttgcac 480
127 tatgctggcg ttgatgttgc agttggaggg atggagcaga gaaaaataca catgttagca 540
129 agggagcttt taccaaaaaa ggttgtttgt attcacaacc ctgtcttaac gggtttggat 600
131 ggagaaggaa agatgagttc ttcaaaaggg aattttatag ctggtgatga ctctccagaa 660
133 gagattaggg ctaagataaa gaaagcatac tgcccagctg gagttggtga aggaaatcca 720
135 ataattggaga tagctaaata cttccttgaa tctcctttaa ccataaaaag gccagaaaaa 780
137 tttggtggag atttgacagt taatagctat gaggagttag agagtttatt taaaaataag 840
139 gaattgcac caatggattt aaaaaatgct gtagctgaag aacttataaa gatttttagag 900
141 ccaattagaa agagatta 918

```

144 &lt;210&gt; SEQ ID NO: 4

145 &lt;211&gt; LENGTH: 918

146 &lt;212&gt; TYPE: DNA

147 &lt;213&gt; ORGANISM: artificial

149 &lt;220&gt; FEATURE:

150 &lt;223&gt; OTHER INFORMATION: tRNA synthetase based on M. jannaschii tyrosyl-tRNA synthetase

152 &lt;400&gt; SEQUENCE: 4

```

153 atggacgaat ttgaaatgat aaagagaaac acatctgaaa ttatcagcga ggaagagtta 60
155 agagagggttt taaaaaaaga tgaaaaatct gctcatatag gttttgaacc aagtggtaaa 120
157 atacatttag ggcattatct ccaaataaaa aagatgattg atttcaaaa tgctggattt 180
159 gatataatta tattgttggc tgatttacac gcctatttaa accagaaagg agagttggat 240
161 gagattagaa aaataggaga ttataacaaa aaagtttttg aagcaatggg gttaaaggca 300
163 aaatatgttt atggaagtga gttccagctt gataaggatt atacactgaa tgtctataga 360
165 ttggctttta aaactacctt aaaaagagca agaaggagta tggaacttat agcaagagag 420

```

## RAW SEQUENCE LISTING

DATE: 11/08/2005

PATENT APPLICATION: US/10/530,421

TIME: 12:07:48

Input Set : A:\54A-000170PC.ST251.txt

Output Set: N:\CRF4\11082005\J530421.raw

```

167 gatgaaaatc caaaggttgc tgaagttatc tatccaataa tgcaggttaa tcggccgcac 480
169 tatcctggcg ttgatgttgc agttggaggg atggagcaga gaaaaataca catgttagca 540
171 agggagcttt taccaaaaaa ggttggttgc attcacaacc ctgtcttaac gggtttggat 600
173 ggagaaggaa agatgagttc ttcaaaaggg aattttatag ctgttgatga ctctccagaa 660
175 gagattaggg ctaagataaa gaaagcatac tgcccagctg gagttgttga aggaaatcca 720
177 ataattggaga tagctaaata cttccttgaa taccctttaa ccataaaaag gccagaaaaa 780
179 tttggtggag atttgacagt taatagctat gaggagttag agagtttatt taaaaataag 840
181 gaattgcac ccaatggattt aaaaaatgct gtagctgaag aacttataaa gatttttagag 900
183 ccaattagaa agagatta 918

```

186 &lt;210&gt; SEQ ID NO: 5

187 &lt;211&gt; LENGTH: 918

188 &lt;212&gt; TYPE: DNA

189 &lt;213&gt; ORGANISM: artificial

191 &lt;220&gt; FEATURE:

192 <223> OTHER INFORMATION: tRNA synthetase based on M. jannaschii tyrosyl-tRNA synthetase

194 &lt;400&gt; SEQUENCE: 5

```

195 atggacgaat ttgaaatgat aaagagaaac acatctgaaa ttatcagcga ggaagagtta 60
197 agagaggttt taaaaaaaga tgaaaaatct gcttatatag gttttgaacc aagtggtaaa 120
199 atacatttag ggcattatct ccaaataaaa aagatgattg atttacaaa tgctggattt 180
201 gatataatta tattgttggc tgattttacac gcctatttaa accagaaagg agagttggat 240
203 gagattagaa aaataggaga ttataacaaa aaagtttttg aagcaatggg gttaaaggca 300
205 aaatatgttt atggaagtcc ttccagctt gataaggatt atacactgaa tgtctataga 360
207 ttggctttta aaactacctt aaaaagagca agaaggagta tggaaactat agcaagagag 420
209 gatgaaaatc caaaggttgc tgaagttatc tatccaataa tgcaggttaa tcagagtcac 480
211 tatgatggcg ttgatgttgc agttggaggg atggagcaga gaaaaataca catgttagca 540
213 agggagcttt taccaaaaaa ggttggttgc attcacaacc ctgtcttaac gggtttggat 600
215 ggagaaggaa agatgagttc ttcaaaaggg aattttatag ctgttgatga ctctccagaa 660
217 gagattaggg ctaagataaa gaaagcatac tgcccagctg gagttgttga aggaaatcca 720
219 ataattggaga tagctaaata cttccttgaa taccctttaa ccataaaaag gccagaaaaa 780
221 tttggtggag atttgacagt taatagctat gaggagttag agagtttatt taaaaataag 840
223 gaattgcac ccaatggattt aaaaaatgct gtagctgaag aacttataaa gatttttagag 900
225 ccaattagaa agagatta 918

```

228 &lt;210&gt; SEQ ID NO: 6

229 &lt;211&gt; LENGTH: 918

230 &lt;212&gt; TYPE: DNA

231 &lt;213&gt; ORGANISM: artificial

233 &lt;220&gt; FEATURE:

234 <223> OTHER INFORMATION: tRNA synthetase based on M. jannaschii tyrosyl-tRNA synthetase

236 &lt;400&gt; SEQUENCE: 6

```

237 atggacgaat ttgaaatgat aaagagaaac acatctgaaa ttatcagcga ggaagagtta 60
239 agagaggttt taaaaaaaga tgaaaaatct gcttcgatag gttttgaacc aagtggtaaa 120
241 atacatttag ggcattatct ccaaataaaa aagatgattg atttacaaa tgctggattt 180
243 gatataatta tattgttggc tgattttacac gcctatttaa accagaaagg agagttggat 240
245 gagattagaa aaataggaga ttataacaaa aaagtttttg aagcaatggg gttaaaggca 300
247 aaatatgttt atggaagtac gttccagctt gataaggatt atacactgaa tgtctataga 360
249 ttggctttta aaactacctt aaaaagagca agaaggagta tggaaactat agcaagagag 420
251 gatgaaaatc caaaggttgc tgaagttatc tatccaataa tgcaggttaa tacgtatcat 480
253 tatgctggcg ttgatgttgc agttggaggg atggagcaga gaaaaataca catgttagca 540
255 agggagcttt taccaaaaaa ggttggttgc attcacaacc ctgtcttaac gggtttggat 600

```

## RAW SEQUENCE LISTING

DATE: 11/08/2005

PATENT APPLICATION: US/10/530,421

TIME: 12:07:48

Input Set : A:\54A-000170PC.ST251.txt

Output Set: N:\CRF4\11082005\J530421.raw

```

257 ggagaaggaa agatgagttc ttcaaaaggg aattttatag ctggtgatga ctctccagaa 660
259 gagattaggg ctaagataaa gaaagcatac tgcccagctg gagttggtga aggaaatcca 720
261 ataatggaga tagctaaata cttccttgaa taccctttaa ccataaaaag gccagaaaaa 780
263 tttggtggag atttgacagt taatagctat gaggagttag agagtttatt taaaaataag 840
265 gaattgcac caatggattt aaaaaatgct gtagctgaag aacttataaa gatttttagag 900
267 ccaattagaa agagatta 918

```

270 &lt;210&gt; SEQ ID NO: 7

271 &lt;211&gt; LENGTH: 918

272 &lt;212&gt; TYPE: DNA

273 &lt;213&gt; ORGANISM: artificial

275 &lt;220&gt; FEATURE:

276 <223> OTHER INFORMATION: tRNA synthetase based on M. jannaschii tyrosyl-tRNA synthetase

278 &lt;400&gt; SEQUENCE: 7

```

279 atggacgaat ttgaaatgat aaagagaaac acatctgaaa ttatcagcga ggaagagtta 60
281 agagaggttt taaaaaaaga tgaaaaatct gctcctatag gttttgaacc aagtggtaaa 120
283 atacatttag ggcattatct ccaaataaaa aagatgattg atttacaaaa tgctggattt 180
285 gatataatta tattgttggc tgatttacac gcctatttta accagaaagg agagttggat 240
287 gagattagaa aaataggaga ttataacaaa aaagtttttg aagcaatggg gttaaaggca 300
289 aaatatgttt atggaagtat gttccagctt gataaggatt atacactgaa tgtctataga 360
291 ttggctttta aaactacctt aaaaagagca agaaggagta tggaaacttat agcaagagag 420
293 gatgaaaatc caaagggttc tgaagttatc tatccaataa tgcagggttaa taatacgcat 480
295 tatgggggag ttgatgttgc agttggaggg atggagcaga gaaaaataca catgttagca 540
297 agggagcttt taccaaaaaa ggttggttgt attcacaacc ctgtcttaac ggggttggat 600
299 ggagaaggaa agatgagttc ttcaaaaggg aattttatag ctggtgatga ctctccagaa 660
301 gagattaggg ctaagataaa gaaagcatac tgcccagctg gagttggtga aggaaatcca 720
303 ataatggaga tagctaaata cttccttgaa taccctttaa ccataaaaag gccagaaaaa 780
305 tttggtggag atttgacagt taatagctat gaggagttag agagtttatt taaaaataag 840
307 gaattgcac caatggattt aaaaaatgct gtagctgaag aacttataaa gatttttagag 900
309 ccaattagaa agagatta 918

```

312 &lt;210&gt; SEQ ID NO: 8

313 &lt;211&gt; LENGTH: 918

314 &lt;212&gt; TYPE: DNA

315 &lt;213&gt; ORGANISM: artificial

317 &lt;220&gt; FEATURE:

318 <223> OTHER INFORMATION: tRNA synthetase based on M. jannaschii tyrosyl-tRNA synthetase

320 &lt;400&gt; SEQUENCE: 8

```

321 atggacgaat ttgaaatgat aaagagaaac acatctgaaa ttatcagcga ggaagagtta 60
323 agagaggttt taaaaaaaga tgaaaaatct gctacgatag gttttgaacc aagtggtaaa 120
325 atacatttag ggcattatct ccaaataaaa aagatgattg atttacaaaa tgctggattt 180
327 gatataatta tattgttggc tgatttacac gcctatttta accagaaagg agagttggat 240
329 gagattagaa aaataggaga ttataacaaa aaagtttttg aagcaatggg gttaaaggca 300
331 aaatatgttt atggaagtca tttccagctt gataaggatt atacactgaa tgtctataga 360
333 ttggctttta aaactacctt aaaaagagca agaaggagta tggaaacttat agcaagagag 420
335 gatgaaaatc caaagggttc tgaagttatc tatccaataa tgcagggttaa tcagactcat 480
337 tatgagggag ttgatgttgc agttggaggg atggagcaga gaaaaataca catgttagca 540
339 agggagcttt taccaaaaaa ggttggttgt attcacaacc ctgtcttaac ggggttggat 600
341 ggagaaggaa agatgagttc ttcaaaaggg aattttatag ctggtgatga ctctccagaa 660
343 gagattaggg ctaagataaa gaaagcatac tgcccagctg gagttggtga aggaaatcca 720
345 ataatggaga tagctaaata cttccttgaa taccctttaa ccataaaaag gccagaaaaa 780

```

## RAW SEQUENCE LISTING

DATE: 11/08/2005

PATENT APPLICATION: US/10/530,421

TIME: 12:07:48

Input Set : A:\54A-000170PC.ST251.txt

Output Set: N:\CRF4\11082005\J530421.raw

```

347 tttggtggag atttgacagt taatagctat gaggagttag agagtttatt taaaaataag      840
349 gaattgcata caatggattt aaaaaatgct gtagctgaag aacttataaa gatttttagag      900
351 ccaattagaa agagatta                                     918
354 <210> SEQ ID NO: 9
355 <211> LENGTH: 918
356 <212> TYPE: DNA
357 <213> ORGANISM: artificial
359 <220> FEATURE:
360 <223> OTHER INFORMATION: tRNA synthetase based on M. jannaschii tyrosyl-tRNA
synthetase
362 <400> SEQUENCE: 9
363 atggacgaat ttgaaatgat aaagagaaac acatctgaaa ttatcagcga ggaagagtta      60
365 agagagggttt taaaaaaaga tgaaaaatct gctcatatag gttttgaacc aagtggtaaa      120
367 atacatttag ggcattatct ccaaataaaa aagatgattg atttacaaaa tgctggattt      180
369 gatataatta tattgttggc tgatttacac gcctatttaa accagaaagg agagttggat      240
371 gagattagaa aaataggaga ttataacaaa aaagtttttg aagcaatggg gttaaaggca      300
373 aaatatgttt atggaagtaa gttccagctt gataaggatt atacactgaa tgtctataga      360
375 ttggctttta aaactacctt aaaaagagca agaaggagta tggaaacttat agcaagagag      420
377 gatgaaaatc caaaggttgc tgaagttatc tatccaataa tgcagggttaa tccgtgtcat      480
379 tatcatggcg ttgatgttgc agttggaggg atggagcaga gaaaaataca catgttagca      540
381 agggagcttt taccaaaaaa ggttgtttgt attcacaacc ctgtcttaac gggtttggat      600
383 ggagaaggaa agatgagttc ttcaaaaggg aattttatag ctgttgatga ctctccagaa      660
385 gagattaggg ctaagataaa gaaagcatat tgcccagctg gagttgttga aggaaatcca      720
387 ataatggaga tagctaaata cttccttgaa tatcctttta ccataaaaag gccagaaaaa      780
389 tttggtggag atttgacagt taatagctat gaggagttag agagtttatt taaaaataag      840
391 gaattgcata caatggattt aaaaaatgct gtagctgaag aacttataaa gatttttagag      900
393 ccaattagaa agagatta                                     918
396 <210> SEQ ID NO: 10
397 <211> LENGTH: 918
398 <212> TYPE: DNA
399 <213> ORGANISM: artificial
401 <220> FEATURE:
402 <223> OTHER INFORMATION: tRNA synthetase based on M. jannaschii tyrosyl-tRNA
synthetase
404 <400> SEQUENCE: 10
405 atggacgaat ttgaaatgat aaagagaaac acatctgaaa ttatcagcga ggaagagtta      60
407 agagagggttt taaaaaaaga tgaaaaatct gctgctatag gttttgaacc aagtggtaaa      120
409 atacatttag ggcattatct ccaaataaaa aagatgattg atttacaaaa tgctggattt      180
411 gatataatta tattgttggc tgatttacac gcctatttaa accagaaagg agagttggat      240
413 gagattagaa aaataggaga ttataacaaa aaagtttttg aagcaatggg gttaaaggca      300
415 aaatatgttt atggaagtcg gttccagctt gataaggatt atacactgaa tgtctataga      360
417 ttggctttta aaactacctt aaaaagagca agaaggagta tggaaacttat agcaagagag      420
419 gatgaaaatc caaaggttgc tgaagttatc tatccaataa tgcagggttaa tgtgattcat      480
421 tatgatggcg ttgatgttgc agttggaggg atggagcaga gaaaaataca catgttagca      540
423 agggagcttt taccaaaaaa ggttgtttgt attcacaacc ctgtcttaac gggtttggat      600
425 ggagaaggaa agatgagttc ttcaaaaggg aattttatag ctgttgatga ctctccagaa      660
427 gagattaggg ctaagataaa gaaagcatat tgcccagctg gagttgttga aggaaatcca      720
429 ataatggaga tagctaaata cttccttgaa tatcctttta ccataaaaag gccagaaaaa      780
431 tttggtggag atttgacagt taatagctat gaggagttag agagtttatt taaaaataag      840
433 gaattgcata caatggattt aaaaaatgct gtagctgaag aacttataaa gatttttagag      900
435 ccaattagaa agagatta                                     918

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/530,421

DATE: 11/08/2005  
TIME: 12:07:49

Input Set : A:\54A-000170PC.ST251.txt  
Output Set: N:\CRF4\11082005\J530421.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:22; Xaa Pos. 7

Invalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,22,23,24



**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/530,421

DATE: 11/08/2005

TIME: 12:07:49

Input Set : A:\54A-000170PC.ST251.txt

Output Set: N:\CRF4\11082005\J530421.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application Number

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:1030 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22 after pos.:0